	2023/2024
Bioinformatics	
Code: 101951 ECTS Credits: 6	

Degree	Type	Year	Semester
2500890 Genetics	OB	3	1

Contact

Name: Sonia Casillas Viladerrams

Email: sonia.casillas@uab.cat

Teaching groups languages

You can check it through this link. To consult the language you will need to enter the CODE of the subject. Please note that this information is provisional until 30 November 2023.

Teachers

Antoni Barbadilla Prados

Marta Coronado Zamora

Oscar Conchillo Solé

Raquel Egea Sanchez

Prerequisites

It is convenient for the student to review the basic concepts acquired in the subjects of the second grade:

- The module of Fundamentals of programming within the subject Instrumental Techniques
- Molecular Genetics of Eukaryotes
- Molecular Genetics of Prokaryotes
- Population Genetics

Also, this subject is complemented by the Genomics, proteomics and interactomics course that is done during the same semester, so taking both together is recommended.

Knowledge of English is highly recommended for reading information.

A basic level of any programming language (Python preferably) will be essential to be able to follow the sessions, both theoretical and practical, and complete the proposed activities.

Objectives and Contextualisation

Bioinformatics -the treatment, computer analysis and the interpretation and modeling of molecular and genomic data- has acquired a fundamental role in genetics nowadays. The subject taught during this course constitutes an introductory vision to bioinformatics. The main objectives are:

- To provide students with the basic bioinformatics knowledge that will allow them both to use a variety of tools to perform information searches in the databases and to approach the computational analysis of sequences of nucleic acids, proteins and genomes
- To apply scripting for the automation of bioinformatic tasks that have been previously studied in class, in order to develop computational thinking
- To give a perspective of the potential of this discipline both in the research and professional shperes

Competences

- Act with ethical responsibility and respect for fundamental rights and duties, diversity and democratic values.
- Be able to analyse and synthesise.
- Be able to communicate effectively, orally and in writing.
- Describe and identify the structural and functional characteristics of nucleic acids and proteins including their different organisational levels.
- Describe the organisation, evolution, inter-individual variation and expression of the human genome.
- Develop self-directed learning.
- Know and apply the omic' tools of genomics, transcriptomics and proteomics.
- Make changes to methods and processes in the area of knowledge in order to provide innovative responses to society's needs and demands.
- Perceive the strategic, industrial and economic importance of genetics and genomics to life sciences, health and society.
- Reason critically.
- Take account of social, economic and environmental impacts when operating within one's own area of knowledge.
- Take sex- or gender-based inequalities into consideration when operating within one's own area of knowledge.
- Use and interpret data sources on the genomes and macromolecules of any species and understand the basics of bioinformatics analysis to establish the corresponding relations between structure, function and evolution.
- Use and manage bibliographic information or computer or Internet resources in the field of study, in one's own languages and in English.

Learning Outcomes

1. Act with ethical responsibility and respect for fundamental rights and duties, diversity and democratic values.
2. Be able to analyse and synthesise.
3. Be able to communicate effectively, orally and in writing.
4. Defend the relevance of progress in the generation and interpretation of data on a genomic scale for our understanding and technological manipulation of organisms.
5. Develop self-directed learning.
6. Explain and apply the methods for the analysis and annotation of genomes.
7. List and explain the content of bioinformatics databases and perform searches for information.
8. Make changes to methods and processes in the area of knowledge in order to provide innovative responses to society's needs and demands.
9. Reason critically.
10. Take account of social, economic and environmental impacts when operating within one's own area of knowledge.

11. Take sex- or gender-based inequalities into consideration when operating within one's own area of knowledge.
12. Use and interpret the results of bioinformatics applications in the molecular analysis of sequences.
13. Use and manage bibliographic information or computer or Internet resources in the field of study, in one's own languages and in English.
14. Use bioinformatics techniques and tools to describe and analyse the human genome.
15. Use the techniques, tools and methodologies used to describe, analyse and interpret the enormous amounts of data produced by high performance technologies.

Content

Topic 1. Introduction to bioinformatics. Bioinformatics in genomic research. The three corners of bioinformatic analysis. The bioinformatic challenges of data integration.

Topic 2. Databases of interest for biosciences. Bibliographic databases. Construction of search strategies. Molecular databases. Search engines.

Topic 3. Sequence formats. Extraction of sequences from the primary databases. Sequence Formats. Tools for conversion between formats.

Topic 4. Sequence statistics. Sequence visualization, editing and analysis tools. Sequence statistics.

Topic 5. Sequence alignment. Comparison of sequences. Local and global alignment for dynamic programming. Scoring matrices: identity and similarity. Penalties for gaps. Dot-plots.

Topic 6. Similarity searches. Heuristic algorithms for similarity searching against databases: BLAST. Applications of BLAST. Search for remote counterparts using PSI-Blast. Tools for the automation of similarity searches.

Topic 7. Multiple alignment. Multiple sequence alignment. Progressive alignments. Distance matrices. The CLUSTAL package. Applications of multiple alignment. Tools for the automation of multiple alignment. Matrices of position (weight) and LOGOS.

Topic 8. Molecular phylogenetic reconstruction. The molecular phylogeny. Methods of phylogenetic inference. Tools for phylogenetic reconstruction. Examples of phylogenetic reconstruction.

Topic 9. Comparative and functional genomics. Different levels of comparison. Synthenic blocks. Rearrangement hotspots. Comparative view using genome browsers.

Topic 10. Genome assembly. Algorithms for mapping reads against reference genomes. Formats and tools.

Topic 11. Genetic variability. Description of genetic variability from VCF files. Calculation of different measures of variability and tests of neutrality. Databases of SNPs and phenotypes.

Methodology

The teaching methodology includes two types of differentiated activities: classes of theory and practical classes in the computer room. Learning will also have individual and/or collective tutoring of students that will support them to solve specific questions or that they require because of their complexity or difficulty.

Theoretical classes

Classes to transmit the basic concepts and information necessary to develop self-learning. Promotion of active participation of students.

Computer classroom practices

These practices will be organized based on problems posed by teachers that will need to be solved using the different tools and bioinformatics analyzes.

Seminars and/or conferences

In the last sessions of the course there will be seminars and/or conferences of researchers in the field of bioinformatics.

Tutorials

Individuals or small groups for the resolution of doubts related to the subject. This type of activity can be done at the request of the students.

Annotation: Within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

Activities

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Computer room practices	27	1.08	4, 5, 6, 7, 9, 3, 2, 13, 12, 14, 15
Seminars	3	0.12	4, 6, 7, 9, 2, 13, 12, 14, 15
Theory	25	1	4, 6, 7, 9, 2, 13, 12, 14, 15
Type: Supervised			
Tutoring	4	0.16	4, 5, 6, 7, 9, 3, 2, 13, 12, 14, 15
Type: Autonomous			
Bibliography reading	2	0.08	4, 5, 9, 3, 2, 13
Computational problem solving	40	1.6	4, 5, 6, 7, 9, 2, 13, 12, 14, 15
Elaboration of works	25	1	4, 5, 6, 7, 9, 3, 2, 13, 12, 14, 15
Study	24	0.96	4, 5, 6, 7, 9, 3, 2, 13, 12, 14, 15

Assessment

The evaluation system is organized in five main activities and there will be, in addition, a recovery exam and an optional final improvement activity. The details of the activities are:

Main evaluation activities

- Partial evaluations (1 and 2). Overall Weight 60%

Partial assessments are combined tests that can consist of multiple-choice questions, written answers and problem solving. These tests will evaluate all the contents taught in the subject in the form of a theoretical-practical test.

The weight of each one of the assessments will be 30%.

In order to pass these evaluations, a minimum grade of 4.0 must be obtained in each of them.

- Computational problems (Rosalind). Overall Weight 25%

Throughout the course, professors will present computational problems or issues related to the subject matter that students will have to solve in the form of evaluations.

The continued nature of this assessment means that the subject can not be evaluated unless there is a minimum participation of 50% of the proposed tests.

- ADAPT YOURSELF-Bioinformatics (integrative work). Overall Weight 15%

This activity is linked to the ADAPT YOURSELF teaching innovation project and gives the option to participate in a learning and service experience in the second semester, recognized with 2 ECTS (https://adaptat.omicsuab.org).

The work will consist of solving a problem that will contain questions related to the different thematic blocks presented during the classes of theory and practices.

These works will be carried out in groups of 3-5 students and are of an autonomous nature. Professors will tutor the work and solve doubtsand questions of approach.

The students will deliver the work following the content guidelines, presentation and deadlines.

This activity is compulsory and non-participation will entail non-evaluation of the subject.

- Computer classroom practices. They can subtract up to 1 point from the final grade

This activity is compulsory and the absence without justification or non-profit of the activity may substract up to 1 point from the final grade of the subject.
The continuous nature of this assessment means that the subject can not be evaluated unless thereis a minimum participation in 50% of the proposed sessions.

- Optional problems or the student's own creations, and active participation in class/forum. They can add up to 1 point on the final grade

Throughout the course the professors will propose complex computational problems that will be optional. Their resolution may add up to 1 point on the final grade of the subject.

Retake exam

To be eligible for the retake process, students should have been previously evaluated in a set of activities equaling at least two thirds of the final score of the course.

The evaluations 1 and/or 2 can be retaken individually and the grade, if >=4, will make an average with those passed in the partial exams. In the event that the grade of the retake exam does not reach a minimum of 4.0, it cannot make an average and the student does not pass the subject.

Continuous assessment, integrative work and attendance to the computer classroom practices, due to its continued nature, cannot be retaken.

Improvement of the final qualification

Students who have passed the assessments 1 and 2 and want to improve their final qualification, may opt for a final test. This test will include all of the subject matter. It is not possible to improve the note through works or other types of activities.

The degree of difficulty of this test will correspond to the objective of the test and, therefore, may be superior to that of the partial evaluations.

The student who submits to this test waives the previous qualifications and therefore, the grade of this improvement test will be the one that will prevail in the final grade even though it islowerthan that obtained by partial.

In no case will the students that take this test opt for the qualification of *Matrícula d'Honor*.

Formula of weighting of the final note

Final note = [(Evaluation 1 x 0.3) +(Evaluation 2 x 0.3) + (Continuous assessment x 0.25) + (Integrative work x 0.15)]

Passed

The subject is considered passed if the final grade is > = 5.0

Not evaluable

The student will be graded as "No Avaluable" if the weighthin of all conducted evaluation activities is less than 67% of the final score.

Unique assessment

The unique assessment consists of a single synthesis test in which the contents of the entire theory and practical program of the subject will be evaluated in the form of a theoretical-practical test. The test will consist of multiple-choice questions, written answers and problem solving. The grade obtained in this synthesis test will account for 60%of the final grade of the subject. In order to pass this synthesis test, a minimum grade of 4.0 must be obtained.

The assessment of Computational Problems (Rosalind) and ADAPT YOURSELF-Bioinformatics (integrative work) will follow the same process as the continuous assessment. The grade obtained will account for 25% and 15% of the final grade of the subject, respectively. Students taking the single assessmentcan hand in these two pieces of evidence together on the same day as the one set for the synthesis test.

The single assessment test will coincide with the same date fixed in the calendar for the last partial assessment test and the same recovery system will be applied as for the conventional assessment.

Students who take the single assessment must do the Computer classroom practices (PLAB) in face-to-face sessions and under the same conditions as in the conventional assessment.

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Computer room practices	Can substract up to 1 point from the final grade	0	0	1, 11, 10, 4, 5, 6, 8, 7, 9, 3, 2, 13, 12, 14, 15
Continuous assessment	25	0	0	4, 5, 6, 7, 9, 3, 2, 13, 12, 14, 15
Integrative work	15	0	0	1, 11, 10, 4, 5, 6, 8, 7, 9, 3, 2, 13, 12, 14, 15
Optional problems or the student's own creation	Can add up to 1 point on the final grade	0	0	1, 11, 10, 4, 5, 6, 8, 7, 9, 3, 2, 13, 12, 14, 15
Partial exam 1 (theoretical-practical contents)	30	0	0	4, 5, 6, 7, 9, 3, 2, 13, 12, 14, 15

Partial exam 2 (theoretical-practical contents)	30	0	0	4, 5, 6, 7, 9, 3, 2, 13, 12, 14, 15

Bibliography

Books

- Attwood, TK. i Parry-Smith, J. 1999. Introduction to Biocomputing. Longman, UK. (*Introducción a la Bioinformática. Attwood y Parry-Smith. 2002. Pearson Educación, S.A.*)
- Compeau, P i Pevzner, P. 2018. Bioinformatics Algorithms. An active learning approach. 3rd Edition. Active Learning Publishers LLC. (read online for free: https://www.bioinformaticsalgorithms.org/)
- Cristianini, N. Y M. W. Hahn. 2007. Introduction to Computational Genomics. A case studies approach. Cambridge Univ. Press. (read online: https://eembdersler.files.wordpress.com/2012/02/introduction-to-computational-genomics-a-case-studies-a)
- Pevzer, P & R Shamir. 2011. Bioinformatics for biologists. Cambridge Univ. Press. (read online: http://docshare04.docshare.tips/files/18157/181579469.pdf)
- Samuelsson, T. 2012. Genomics and Bioinformatics. Cambridge Univ. Press.
- Scherer, S. 2008. A short guide to the Human Genome. Cold Spring Harbor Laboratory Press.
- Xiong, J. 2006. Essential bioinformatics. Cambridge Univ. Press.

MOOCS and online resources:

- Biology meets Programming: Bioinformatics for Beginners (course by Compeau&Pevzner): https://www.coursera.org/learn/bioinformatics (easy! for everyone!)
- Guions de pràctiques de l'assignatura: https://stepik.org/course/1171/syllabus
- Journey to the Frontier of Computational Biology (Specialization by Compeau&Pevzner): https://www.coursera.org/specializations/computational-biology (hard! for those who want a big challenge!)
- Problemes computacionals a Rosalind: http://rosalind.info/problems/locations/
- Web associada al llibre "Compeau, P i Pevzner, P. 2018. Bioinformatics Algorithms. An active learning approach. 3rd Edition. Active Learning Publishers LLC": https://www.bioinformaticsalgorithms.org/

Software

Basic software

- Problemes computacionals a Rosalind: http://rosalind.info/problems/locations
- Llenguatge de programació Python (i BioPython)
- Anaconda
- Jupyter notebook

Databases

- NCBI-Entrez Gene http://www.ncbi.nlm.nih.gov/gene
- NCBI-RefSeq http://www.ncbi.nlm.nih.gov/refseq/
- NCBI-Genome https://www.ncbi.nlm.nih.gov/genome/
- NCBI-Genbank https://www.ncbi.nlm.nih.gov/genbank/
- OMIM https://www.omim.org/
- GeneCards https://www.genecards.org/
- Geneid https://genome.crg.es/geneid.html
- Genscan http://hollywood.mit.edu/GENSCAN.html
- Pubmed http://www.ncbi.nlm.nih.gov/pubmed
- dbSNP https://www.ncbi.nlm.nih.gov/snp/
- ENSEMBL http://www.ensembl.org/

- PopHuman https://pophuman.uab.cat/
- PopHumanScan https://pophumanscan.uab.cat/
- EVA https://www.ebi.ac.uk/eva/?Home
- GWAS Catalogue https://www.ebi.ac.uk/gwas/docs/about
- ClinVar https://www.ncbi.nlm.nih.gov/clinvar/intro/
- DisGeNET https://www.disgenet.org/help
- Open Targets Platform https://platform.opentargets.org/
- RegulomeDB https://regulomedb.org/
- Human Genome Dating https://human.genome.dating/

Analysis tools

- EMBOSS http://www.bioinformatics.nl/emboss-explorer/ , http://www.ebi.ac.uk/Tools/emboss/
- Galaxy https://usegalaxy.org/ (i moltes eines integrades en aquesta plataforma)
- BLAST http://blast.ncbi.nlm.nih.gov/Blast.cgi
- Transeq https://www.ebi.ac.uk/Tools/st/emboss_transeq/
- PSA https://www.ebi.ac.uk/Tools/psa/
- MSA https://www.ebi.ac.uk/Tools/msa/
- Clustal Omega https://www.ebi.ac.uk/Tools/msa/clustalo/
- Clustal W http://www.clustal.org/download/current/clustalw-2.1-win.msi
- MEGA http://www.megasoftware.net/